8

Realism about what?

It is tempting to dismiss worries about the underdetermination of theories as purely philosophical doubts. Of course we cannot prove beyond all doubt that electrons exist, but then we cannot prove beyond all doubt that the Sun will rise tomorrow, that all metals expand when heated, or that everyday objects such as tables are still there when we aren't perceiving them. Hence, assuming that the mere possibility of error is not sufficient seriously to undermine ordinary claims to know, neither does it threaten scientific realism. Indeed, as we saw in the previous chapter, the very most that van Fraassen can claim to have shown is that it is not irrational for us to withhold belief in unobservables, but neither is it irrational to believe in them, and this is much weaker than traditional forms of scepticism about science. In this chapter, we will consider arguments for various kinds of antirealism, which are motivated by careful scrutiny of the practice and history of science, rather than by epistemological scruples. In different ways, facts about real science raise the question 'what should we be realists about?'.

8.1 Theory change

Perhaps the most compelling argument against scientific realism is the notorious 'pessimistic meta-induction', which in recent years has been championed by Larry Laudan (1981). This is a form of 'argument from theory change', and differs from the underdetermination argument by appealing to history, rather than to theoretical

possibilities and baroque mathematical constructions. Recall the (global IBE) argument of 7.2.2; roughly, there is an explanatory connection between the empirical success of scientific theories and their truth, in virtue of which scientific realism offers the only, or at least the best, explanation of the progress of science. Laudan turns this argument around and argues that we have positive reason, by induction, for not believing in the existence of the theoretical entities described by our best current theories. Like the global defence of realism, this argument is at the 'meta-level' because it is premised on consideration of science as a whole and its evolution in time. Laudan produced a list of now abandoned theories that once enjoyed predictive and explanatory success. Many of these theories featured theoretical terms, such as 'phlogiston', 'ether' and 'caloric', which were supposed to refer to various kinds of unobservables but which, according to modern science, fail to refer to anything at all (just as the terms 'unicorn', 'dragon' and 'leprechaun' don't refer to anything). If this is so then, rather than being justified in believing our best current theories are true, we have inductive grounds for the pessimistic conclusion that they too are likely to be replaced by successors which will show them to be false. Consequently, we have good reason to believe that terms like 'electron' don't refer to unobservable particles of matter after all. Hence, whatever the merits of IBE and whether or not constructive empiricism is ultimately defensible, scientific realism cannot be the best explanation of the success of science because it is not even empirically adequate. Notice that this is an argument for *atheism* about unobservable entities – believing that electrons and the like don't exist – not van Fraassen's type of *agnosticism*.

8.1.1 Approximate truth

For a while, in the eighteenth and nineteenth centuries, it seemed as if Newtonian mechanics might offer a completely true account of the behaviour of matter. It was widely thought that light was composed of particles and therefore that Newton's laws of motion must underlie optics too. Indeed, when Maxwell's electromagnetism became widely accepted many physicists thought that this in no way impugned the status of Newtonian theory because they assumed that

somehow the former would eventually be reduced to the latter. The confidence in classical physics was such that, at the end of the nineteenth century, some physicists apparently thought that their subject was nearly finished and only a few problems remained to be ironed out. Even the philosophically sophisticated Poincaré, whose knowledge of contemporary science and its history was unrivalled, was convinced that Euclidean geometry and Newtonian mechanics would always reign supreme. However, subsequently, the discovery of radioactivity opened up a whole new domain of inquiry and eventually led to the positing of two new fundamental forces (the strong and weak nuclear forces), where classical physics recognised only gravity and electromagnetism as fundamental. Furthermore, of course, Einstein's theories of special and general relativity forced a major revolution in the understanding of gravity, space and time. Similar revolutions have taken place in chemistry; for example, the adoption of the modern theory of atomic structure; and in biology, for example the acceptance of evolution and then DNA. Hence, nowadays, even those who have great faith in science and never question scientific realism are not so naïve as to think that all the best contemporary scientific theories are wholly true and without fault. There have been enough cases of successful theories being amended in unexpected ways, or of unforeseen new phenomena being discovered, for it to be apparent that even the finest science is subject to revision and correction.

Of course, this is partly what motivated Popper to think that all theories are merely conjectures rather than certitudes, but, as we saw in Chapter 3, he abandoned the idea that we ever have any positive reason to believe theories on the basis of evidence. On the other hand, most scientific realists have a conception of evidential support which entails that the predictive and explanatory success of a theory can provide inductive grounds for believing it to be true. They try and accommodate the lessons of history by claiming that theories are not perfectly but only 'approximately true'. Approximate truth, which is sometimes called *verisimilitude*, is indispensable to contemporary scientific realists, but despite a good deal of work it seems to have eluded a satisfactory, precise characterisation. Popper attempted a formal definition but it famously failed (see Psillos 1999: 261–264). Subsequently, people have developed formal mathematical and

logical theories of approximate truth which they then apply to scientific theories. However, none of these is without its difficulties, and verisimilitude continues to puzzle philosophers.

Even though the idea of approximate truth and the idea of different theories being more or less approximately true than each other are very difficult to explicate, they are arguably needed for an adequate semantics of other types of propositions, not just theoretical scientific ones. For example, if I say that it is midday when it is actually 2 minutes past noon, then my utterance is strictly speaking false, and yet it is undeniably approximately correct. Where we are talking about numbers, the notion of approximate truth is very clear, because we know what it means for one number to be relatively close to another one. In such cases we can even quantify the degree of approximation. However, even when no numbers are involved, a notion of approximate truth may still be intelligible. For example, suppose someone declares that it is raining outside, but in fact it is sleeting or snowing. Clearly, what was said was true in so far as there is precipitation occurring, although false in so far as the form of the precipitation is different from what was stated. Similarly, it is approximately true that the sea is blue although it is usually tinged with green, and so on.

In science, examples of approximation and verisimilitude abound. Every schoolchild learns that the Earth is not flat but spherical, and we picture it as a globe, but of course, the surface of the Earth is very uneven with mountains and valleys, and furthermore it is slightly flattened at the poles. Strictly speaking then, it is not a sphere at all, yet we continue to describe it as such because that is close to the truth in salient respects. Similarly, if modern science is to be believed, theories such as Newtonian mechanics and Maxwell's electrodynamics are false, but scientific realists argue that they were, and are, approximately true, although less so than contemporary theories. Of course, to some extent it is approximately true that the Earth is flat, because it is so big relative to us that for many purposes the degree of curvature is negligible. The danger with the idea of verisimilitude is that it may inherit the permissiveness and relativism of similarity in general. Any two objects are similar in some respect. Take, for example, the Sun and my kettle; they are similar to each other because they are both further from the

centre of the Earth than the bottom of the Pacific ocean is. Judgements of similarity can only be evaluated relative to a pragmatic context. For example, a glass tumbler is more similar to a cup than to a lens relative to function, but more similar to a lens than to a cup relative to constitution. So, if all that realists claim is that the best contemporary scientific theories are similar to the truth, this might be a very weak claim. (Antirealists may argue that the notion of approximate truth is in danger of collapsing onto the notion of approximate empirical adequacy. To say that the predictions of theories for what we can observe are approximately accurate does not amount to a defence of realism.)

Nonetheless, since it is difficult to see how we could get by without some notion of verisimilitude in analysing discourse about the observable world, it would be unreasonable to reject scientific realism, in particular, merely on the grounds that the notion of approximate truth is somewhat vague. In any case, as we shall see, in an attempt to give some content to the claim that a theory is approximately true, many realists have focused on the idea of successful reference. A term can be said *to refer* successfully if there is something, or some things, which are picked out by it. So, for example, 'horse' refers but 'the present king of France' does not. In his challenge to realism, Laudan (1981) argues that a scientific theory cannot be even approximately true if its central theoretical terms do not refer to anything. Hence, in the debate about the extent to which theory change in the history of science motivates scepticism about unobservables, it is often the referential status of abandoned theoretical terms that is in dispute.

8.1.2 Sense and reference

Analytic philosophers often distinguish between the *sense* of a term and its *reference*. Sense is a matter of the ideas and descriptions associated with the term, whereas the reference is the thing or things the term is used to talk about. Sense and reference are, of course, related, but they are not the same. If they were it would be impossible for us to continue to talk about the same thing if we changed our minds about its properties. The sense of the term 'whale' used to include the concept of a fish, whereas this has now been replaced by

the concept of a mammal. Yet 'whale' still refers to the animals that our ancestors enthusiastically hunted. In the case of observable entities, reference can be fixed by pointing to the thing or things in question. This is what philosophers call 'ostensive definition'. This will work even when the sense of the term is at odds with the properties of the thing to which it refers. For instance, someone might say 'Audrey is the one over there talking to Angela'. Suppose that the former woman is in fact talking to someone else, but both parties in the conversation falsely believe her interlocutor to be Angela. Then the description associated with the introduction of the term 'Audrey' will fail to pick out Audrey correctly, yet both people can still successfully refer to her.

The problem is that theoretical terms in science that are alleged by realists to refer to unobservable entities cannot be given an ostensive definition. Instead, it is natural to think that the reference of a term like 'electron' is fixed by the theory of electrons. Therefore, it refers to very small entities that are negatively charged, orbit the nuclei of atoms, have a particular mass, and so on. This amounts to saying that the sense of such terms fixes their reference, and yet, as we saw in Chapter 4, Kuhn pointed out that the sense of many scientific terms such as 'atom', 'electron', 'species' and 'mass' has changed considerably during the course of scientific revolutions. If the reference of theoretical terms is fixed by the whole of the theories in which they feature, then any change in the latter will result in a change in the former. It seems as if we must either accept that the world itself changes when theories change, or that most theoretical terms don't refer to anything.

In response to Kuhn, Putnam (1975b) proposed a radically different account of the meaning of theoretical terms. He pointed out that most of us have no idea how many terms get their reference, but we nonetheless successfully refer to particular kinds of things with them. For example, which things in the world are referred to by the terms 'gold', 'elm' or 'French spaniel' is fixed by a few experts, to whom everyone else defers, yet such terms are part of a common language which we all use despite our ignorance. It is not just that most of us lack an explicit definition, we also often lack any way of distinguishing samples. The experts, on the other hand, have detailed criteria, often requiring subtle empirical tests; for example, in the case of gold

these would include weight, reaction with particular acids, electrical resistance, melting point and so on. Putnam calls this *the division of linguistic labour*.

According to Putnam's theory of meaning for kind terms there are four components. Take the term 'water' for example. First there is its *syntactic marker*, which is that of 'mass noun' (like 'air' or 'wood', as opposed to count nouns, like 'horse' or 'tree'). Then there is the *semantic marker*, which is the idea of a common liquid, and the *stereotype*, which is that water falls from the sky as rain, is drinkable, transparent and so on. Finally, there is the *extension*, which is the actual stuff, the natural kind H_2O, to which 'water' refers. Putnam advocates a 'causal theory of reference' for natural kind terms such as 'water', 'gold' and 'electron', according to which 'water' is whatever causes the experiences that give rise to water talk. Reference is fixed not by the description associated with a term but by the cause that lies behind the term's use. This theory allows for continuity of reference across theory changes. Scientific theories help fix the stereotypes of terms such as 'electron', even though theories about electrons have changed, and hence the meaning of the term has changed, Putnam argues that 'electron' has always referred to whatever causes the phenomena that prompted its introduction, such as, for example, the conduction of electricity by metals. (In the nineteenth century, 'electron' was introduced to refer to the least unit of electric charge; however, physicists now believe that some quarks have a charge that is one third of the magnitude of the charge of electrons.) According to Putnam's theory, the reference of theoretical terms can be very stable indeed; the worry, as we shall see, is that it makes successful reference too easy to achieve.

8.1.3 The pessimistic meta-induction

The meta-induction has the following structure:

(i) There have been many empirically successful theories in the history of science that have subsequently been rejected and whose theoretical terms do not refer according to our best current theories.
(ii) Our best current theories are no different in kind from those

discarded theories and so we have no reason to think they will not ultimately be replaced as well.

(iii) By induction, we have positive reason to expect that our best current theories will be replaced by new theories, according to which some of the central theoretical terms of our best current theories do not refer.

Therefore, we should not believe in the approximate truth or the successful reference of the theoretical terms of our best current theories.

In support of premise (i) Laudan lists the following theories, all of which he claims were empirically successful, but which, if our best contemporary theories are to be believed, have theoretical terms that do not refer:

> the crystalline spheres of ancient and medieval astronomy;
> the humoral theory of medicine;
> the effluvial theory of static electricity;
> the 'catastrophist' geology, with its commitment to a universal (Noachian) deluge;
> the phlogiston theory of chemistry;
> the caloric theory of heat;
> the vibratory theory of heat;
> the vital force theories of physiology;
> the electromagnetic ether;
> the optical ether;
> the theory of circular inertia; and
> theories of spontaneous generation.
>
> (Laudan 1981: 29)

Laudan contends that his list could be extended 'ad nauseam' (Laudan 1981: 29)

However, the first realist strategy for dealing with the meta-induction is to restrict the number of theories that can legitimately be used as the basis for an induction whose conclusion concerns contemporary theories. It would be absurd to claim that we should be realists about any scientific theory, only those with the right features will qualify. One of the most common reactions to the list above is to complain that many of the theories mentioned are nothing like the

best contemporary theories. Hence, although premise (i) above seems undeniable, premise (ii) is not.

(1) Realist responses

(a) Restrict realism to *mature* theories

Most of Laudan's examples have not been taken seriously by realists. They reduce the number that make up the inductive basis, by only considering 'mature' theories from the past as relevant to contemporary science. A science reaches maturity when it satisfies requirements such as coherence with the basic principles of theories in other domains, and possession of a well-entrenched set of basic principles which define the domain of the science and the appropriate methods for it, and limit the sorts of theories that can be proposed. All parts of contemporary physical science incorporate the law of conservation of energy, as well as the basic theory of the structure of matter in terms of elements such as hydrogen, oxygen, carbon and so on. They employ a common system of units such as metre, kilogram, amp, volt, degree, and so on, and they all use concepts such as force, velocity, mass, charge and so on. Furthermore, theories in one domain are routinely deployed as background theories in other domains. Hence, it is arguable that contemporary science has a degree of unification and coherence, as well as mathematical sophistication, that is quite absent in many of theories Laudan cites.

(b) Restrict realism to theories enjoying *novel* predictive success

Many realists have argued that Laudan's notion of empirical success is much too permissive, and so they defend realism against the meta-induction by placing constraints on the kind of empirical success that justifies realism about a theory. After all, the right empirical consequences may be merely written into almost any theory in an ad hoc manner, yet we would not want to count such a theory's empirical success as evidence for its truth, not least because we could then be justified in believing each of two theories that attribute different causal structures to the world to be true. For example, for a time, the wave and particle theories of light both had a good degree of

predictive and explanatory success in relation to known types of phenomena, yet they are ontologically incompatible. Many realists seem to have concluded that a theory needs to have not just empirical success, but to have yielded confirmed *novel* predictions.

The idea of novel predictions was celebrated by Popper (see 3.1). He contrasted the risky predictions of physics with the vague predictions of psychoanalysis, but he also wanted to justify the failure of scientists to abandon Newtonian theory when it was known to be incompatible with certain observations. As we saw in 3.5(5), often various modifications to background assumptions are made to try and accommodate observed facts that would otherwise refute established theories. Popper, and following him Lakatos and others, argued that this course of action is acceptable only when the new theory produces testable consequences other than the results that motivated it. So, for example, the postulation of a new planet to accommodate the observed orbit of a familiar one is legitimate because it ought to be possible to observe the former (or at least its effects on other bodies).

There is a general issue here concerning the confirmation of theories by evidence. Is when a particular piece of evidence was gathered relevant to the degree to which it confirms a hypothesis? *Predictionists* say that only new evidence confirms theories, whereas *explanationists* say that only the explaining of known facts confirms theories. Other philosophers think that both old and new evidence can confirm theories but that new evidence is particularly compelling. Similarly, it is the ability of theories to predict previously unsuspected results that many realists think is particularly miraculous on the supposition that realism is false. To see why, consider the following famous example.

In the early nineteenth century, both the wave and particle theories of light offered competing accounts of familiar optical phenomena. A ray of light directed at a plane mirror at a given angle will be reflected at the same angle (the law of reflection), which is just how a ball behaves when shot at the wall of a billiard table. Hence, the established laws of mechanics, in particular the conservation of momentum, used to explain the behaviour of material objects such as billiard balls, could also be used to the explain the law of reflection in terms of the particle theory of light. On the other hand, other

phenomena were more naturally accommodated by the wave theory. The famous two-slit experiment showed that, in certain circumstances, light passing through two slits and hitting a screen would create a pattern of equally spaced light and dark bands, rather than two illuminated patches opposite the slits and shadow in between. This was explained by wave theorists in terms of light waves overlapping, and so adding to or cancelling each other's intensity at different points in space.

In 1818, Fresnel developed a mathematical theory according to which light consists of waves of a particular kind, namely *transverse* waves. A transverse wave is one that oscillates perpendicularly to its direction of motion, for example a wave on the surface of water. Fresnel's theory made derivations of known effects precise and elegant, but it had a new consequence. It predicted that, in certain circumstances, light that was shone on a completely opaque disk would cast a shadow with a bright white spot in its centre. The phenomenon is known as *conical refraction* and has now been observed many times. However, Fresnel knew nothing of it when he developed his theory, and indeed did not even derive the result himself. This is more striking than the prediction of the existence of an extra planet, because it is a prediction of a completely new and unexpected type of phenomenon. (Another good example – mentioned briefly in 3.1 and again in 6.1.3 – is the prediction of the general theory of relativity that light will take a curved path in the presence of very massive objects such as stars.)

(2) Novelty

If the meta-induction is to be defeated by restricting the theories that can form the inductive base to mature theories that enjoyed novel predictive success, the notion of novelty employed needs to be explained precisely. The most straightforward idea is that of *temporal novelty*. A prediction is temporally novel when it is of something that has not yet been observed. The worry about using this kind of novel predictive success as a criterion for adopting realism about particular theories is that it seems to introduce an element of arbitrariness into which theories are to be believed. When exactly in time someone first observes some phenomenon entailed by a

theory may have nothing to do with how and why the theory was developed. It is surely not relevant to whether a prediction of a theory is novel or not, if it has in fact been confirmed by someone independently who didn't tell anyone about it. (As it turns out, the white spot phenomenon had been observed independently prior to its prediction by Fresnel's theory.) A temporal account of novelty would make the question of whether a result was novel for a theory a matter of mere historical accident and this would undermine the epistemic import that novel success is suppose to have for a particular theory.

It is more plausible to argue that what matters in determining whether a result is novel is whether a particular scientist knew about the result before constructing the theory that predicts it. Call this *epistemic novelty*. The problem with this account of novelty is that, in some cases, the fact that a scientist knew about a result does not seem to undermine the novel status of the result relative to the theory, because the scientist may not have appealed to the former in constructing the latter. For example, many physicists regarded the success of general relativity in accounting for the well-known, previously anomalous orbit of Mercury as highly confirming, because the reasoning that led to the theory appealed to general principles and constraints that had nothing to do with the empirical data about the orbits of planets. The derivation of the correct orbit was not achieved by putting in the right answer by hand.

Worrall suggests realism is only appropriate in the case of: 'theories, designed with one set of data in mind, that have turned out to predict entirely unexpectedly some further general phenomenon' (Worrall 1994: 4). However, in a recent analysis, Leplin rejects this as relativising novelty to the theorist and thereby introducing a psychological and so non-epistemic dimension to novelty (Leplin 1997). Take the case of Fresnel. If we say that the fact that the white spot phenomenon was known about is irrelevant, because Fresnel was not constructing his theory to account for it but it still predicted it, then we seem to be saying that the intentions of a theorist in constructing a theory determine in part whether the success of the theory is to be counted as evidence for its truth. Leplin argues that this undermines the objective character of theory confirmation needed for realism.

This motivates the idea of *use novelty*. A result is use-novel if the scientist did not explicitly build the result into the theory or use it to set the value of some parameter crucial to its derivation. Hence, Leplin proposes two conditions for an observational result *O* to be novel for a theory *T*.

> *Independence condition*: There is a minimally adequate reconstruction of the reasoning leading to *T* that does not cite any qualitative generalization of *O*.
> *Uniqueness condition*: There is some qualitative generalization of *O* that *T* explains and predicts, and of which, at the time that *T* first does so, no alternative theory provides a viable reason to expect instances.
>
> (Leplin 1997: 77)

Leplin explains that a reconstruction of the reasoning that led to a theory is an idealisation of the thought of the theorist responsible for that theory, and is said to be 'adequate' if it motivates proposing it, and 'minimally' so if it is the smallest such chain of reasoning (Leplin 1997: 68–71).

According to the two conditions above, novelty is a complex relation between a theory, a prediction or explanation, the reconstruction of the reasoning that led to it, and all the other theories around at the time, since the latter are required not to offer explanations of the result. It follows that: (a) if we found a dead scientist's revolutionary new theory of physics, but they left no record of what experiments they knew about or what reasoning they employed, such a theory could have no novel success. Hence, no amount of successful prediction of previously unsuspected phenomena would motivate a realist construal of the theory; (b) suppose that we already knew all the phenomena in some domain. In such a case we could never have evidence for the truth rather than the empirical adequacy of any theories that we constructed in this domain, no matter how explanatory, simple, unified or whatever. These consequences are surely contrary to realist thinking. Certainly scientific methodology includes far broader criteria for empirical success, such as providing explanations of previously mysterious phenomena. Indeed, Darwin's theory of evolution and Lyell's theory of uniformitarianism (according to which geological change is the result of ordinary forces acting over

very long periods of time rather than sudden and dramatic catastrophes) were accepted by the scientific community because of their systematising and explanatory power, and in spite of their lack of novel predictive success. Realists also often argue that the unifying power of theories, which may bring about novel predictive success but need not do so, be taken as a reason for being realistic about them.

Furthermore, Leplin admits that his analysis makes novelty temporally indexed (Leplin 1997: 97), but this seems to fall foul of his own claim that 'it is surely unintuitive to make one's epistemology depend on what could, after all, be simple happenstance' (Leplin 1997: 43). The main problem seems to be with the uniqueness condition since it seems to leave too much to chance. For example, suppose a result is novel with respect to some theory, but that another theory comes along soon afterwards which also explains it. According to Leplin's view, a realist commitment to the former theory is warranted but not to the latter, yet it seems that the order of things might have been reversed, so that which theory is believed to be true is to some degree historically contingent. Moreover, truth is being imputed to explain the mystery of the novel success of one theory, but the success of the other theory, which would be novel were its rival not around, is left unexplained. This may leave us wondering whether the explanation of success in terms of truth is really necessary theory for either theory.

This does not mean that novel predictive success is a complete red herring. On the contrary, it does seem that the fact that theories sometimes produce predictions of qualitatively new types of phenomena, which are then subsequently observed, is a strong motivation for some sort of realism about scientific knowledge. However, it is the fact that novel predictive success is possible and happens at all that needs accounting for in terms of a general account of science and the world. On the other hand (like Psillos 1999), Leplin uses whether or not a theory enjoyed novel predictive success as a criterion for whether or not it, in particular, should be believed to be true, which leaves the fate of theories too much to chance.

8.1.4 Counter-examples to the no-miracles argument

Even if many of Laudan's examples are spurious, there are still some examples of theories that were both mature and that enjoyed novel predictive success. Even if there are only one or two such cases, the realist's claim that approximate truth explains empirical success will no longer serve to establish realism. This is because, where a theory is empirically successful, but is not approximately true, we will need some other explanation. If this will do for some theories then it ought to do for all, and then we do not need the realist's preferred explanation that such theories are true. Hence, we do not need to form an inductive argument based on Laudan's list to undermine the no-miracles argument for realism. Laudan's paper was also intended to show that the successful reference of its theoretical terms is *not* a necessary condition for the novel predictive success of a theory (Laudan 1981: 45), in other words, that there are counter-examples to the no-miracles argument:

(i) There are examples of theories that were mature and had novel predictive success, but whose central theoretical terms do not refer according to our best current theories.
(ii) Successful reference of its central theoretical terms is a necessary condition for approximate truth. (Premise (iv) of the meta-induction.)
(iii) There are examples of theories that were mature and had novel predictive success but which are not approximately true. (From (i) and (ii).)
(iv) Approximate truth and successful reference of central theoretical terms is not a necessary condition for the novel-predictive success of scientific theories.
(v) The no-miracles argument is undermined since, if approximate truth and successful reference are not available to be part of the explanation of some theories' novel predictive success, there is no reason to think that the novel predictive success of other theories has to be explained by realism.

Most attention has focused on the examples of the '*ether* theory of light' and the '*caloric* theory of heat'. Whatever the realists' preferred way of characterising novel predictive success, it seems these theories had it, so they need to account for these cases.

(1) Realist responses to the counter-examples

There are two basic responses to the counter-examples.

(I) Develop an account of reference according to which the relevant abandoned theoretical terms refer after all

In 8.1.2 above we saw that realists have used causal theories of reference to account for the continuity of reference for terms such as 'atom' or 'electron', when the theories about atoms and electrons undergo significant changes. The difference with the terms 'ether' and 'caloric' is that they are no longer used in modern science. In the nineteenth century, the ether was usually envisaged as some sort of material, solid or liquid, which permeated all of space. It was thought that light waves had to be waves in some sort of medium and the ether was posited to fulfil this role. Yet if there really is such a medium then we ought to be able to detect the effect of the Earth's motion through it, because light waves emitted perpendicular to the motion of a light source through the ether ought to travel a longer path than light waves emitted in the same direction as the motion of the source through the ether. Various experiments failed to find such an effect. Furthermore, soon after Fresenel's ether theory of light had its great successes, Maxwell developed his theory of the electromagnetic field. Light was now regarded as a wave in the electromagnetic field, which was not considered to be a material substance at all. As a result, the term 'ether' was eventually abandoned completely.

However, the causal theory of reference may be used to defend the claim that the term 'ether' referred after all, but to the electromagnetic field and not to a material medium. If the reference of theoretical terms is to whatever causes the phenomena responsible for the terms' introduction, then since optical phenomena are now believed to be caused by the oscillations in the electromagnetic field, than the latter is what is referred to by the term 'ether'. Similarly, since heat is now believed to be caused by molecular motions, then the term 'caloric' can be thought to have referred all along to these rather than to a material substance. The worry about this is that it may imply that reference of theoretical terms is a trivial matter, since as long as some phenomena prompt the introduction of a term it will

automatically successfully refer to whatever is the relevant cause (or causes). Furthermore, this theory radically disconnects what a theorist is talking about from what they think they are talking about. For example, Aristotle and Newton could be said to be referring to geodesic motion in a curved spacetime when, respectively, they talked about the natural motion of material objects, and the fall of a body under the effect of the gravitational force. We shall return to this issue below.

(II) Restrict realism to those theoretical claims about unobservables that feature in an *essential* way in the derivation of novel predictions

The essence of this strategy is to argue that the parts of theories that have been abandoned were not really involved in the production of novel predictive success. Philip Kitcher says that: 'No sensible realist should ever want to assert that the idle parts of an individual practice, past or present, are justified by the success of the whole' (Kitcher 1993: 142).

Similarly, Psillos argues that history does not undermine a cautious scientific realism that differentiates between the evidential support that accrues to different parts of theories, and only advocates belief in those parts that are essentially involved in the production of novel predictive success. This cautious, rather than an all or nothing, realism would not have recommended belief in the parts of the theories to which Laudan draws attention because, if we separate the components of a theory that *generated* its success from those that did not, we find that the theoretical commitments that were subsequently abandoned are the *idle* ones. On the other hand, argues Psillos: 'the theoretical laws and mechanisms that generated the successes of past theories have been retained in our current scientific image' (Psillos 1999: 108).

Such an argument needs to be accompanied by specific analyses of particular theories that both (a) identify the essential contributors to the success of the theory in question, and (b) show that these were retained in subsequent developments.

Psillos's strategy for defeating the argument from theory change is a combination of (I) and (II). Laudan claims that if current successful

theories are approximately true, then the caloric and ether theories cannot be because their central theoretical terms do not refer (by premise (ii) above). Strategy (I) accepts premise (ii) but Psillos allows that sometimes an overall approximately true theory may fail to refer. He then undercuts Laudan's argument by arguing that:

- Abandoned theoretical terms that do not refer, such as 'caloric', were involved in parts of theories not supported by the evidence at the time, because the empirical success of caloric theories was independent of any hypotheses about the nature of caloric.
- Abandoned terms that were used in parts of theories supported by the evidence at the time do refer after all; 'ether' refers to the electromagnetic field.

Below, I identify some problems with this type of defence of realism and the claims made about 'ether' and 'caloric'. Most importantly, I argue that the notion of 'essential', deployed in (a) above is too vague to support a principled distinction between our epistemic attitudes to different parts of theories. In my view, ultimately the problem with strategy (II) is that it is in danger of becoming ad hoc and dependent on hindsight.

According to Psillos: 'Theoretical constituents which make essential contributions to successes are those that have an indispensable role in their generation. They are those which "really fuel the derivation" ' (Psillos 1999: 110). This means that the hypothesis in question cannot be replaced by an independently motivated, non ad hoc, potentially explanatory alternative. (Remember that the sort of success referred to here is novel predictive success.) Psillos gives as an example of an idle component of a theory Newton's hypothesis that the centre of mass of the universe is at absolute rest. However, within Newton's system this hypothesis cannot be replaced by an alternative that satisfies the above requirements. That the centre of mass of the universe should move with any particular velocity is surely more ad hoc than it being at rest. There is also a sense in which the universe being at rest is simpler than it having a particular velocity, since the latter would raise the further question of what force set it in motion, requiring a new theory to explain its motion. Certainly, any alternative hypotheses that might have been entertained would not be explanatory of anything nor independently motivated. It seems that

this hypothesis does count as an essential contribution to the success of Newton's theory, by Psillos' criteria, and hence that he would have had us be realistic about it. However, the notions of absolute space and absolute rest have no meaning in modern physics, so Psillos' criterion above has accidentally bolstered historically inspired scepticism.

This case is arguably not so serious for the realist for it does not involve a central theoretical term that was essentially deployed, yet which cannot be regarded as referring from the standpoint of later theory. Nonetheless, Psillos is intending to deal with the threat of such examples using this distinction between essential and inessential theoretical constituents, and this example reveals that his definition of this distinction does not, in general, capture only the theoretical hypotheses with which the realist would be happy. Another problem is the ambiguity concerning the type of dependence in question when we ask if a theory's success is dependent on a particular hypothesis. We can understand this as at least either logical/mathematical dependence or causal dependence. So, when we are asked to look at the particular novel empirical success of a theory and decide which are the parts of theory that this success depended on, then we will give different answers according to how we understand dependence.

Furthermore, the realist should be careful here for it is dangerous for realism overall to disconnect the metaphysical hypotheses and background assumptions about supposed entities, such as caloric or the ether, from what are construed as the real successes of the theories in issuing certain predictions. One of the central claims of contemporary realism is that we have to take seriously the involvement of theoretical and metaphysical beliefs in scientific methodology; that is, that we cannot disconnect the success of science from the theoretically informed methods of scientists. This is meant to support realism, for, according to Boyd and others, only realism explains why these extra empirical beliefs are so important. However, Psillos suggests that, after all, we need not take seriously scientists' beliefs about the constitution of the ether or caloric, because the success of the theories floats free of such assumptions. Let us now turn to the examples he discusses.

The case of the ether

Ether theories were successful by any plausible criteria that the realist may concoct and were mature, yet it was widely believed that the ether was a material substance, and there certainly is no such thing permeating space according to Maxwell's successor theory. The hypotheses about the material nature of the ether were no presuppositional or idle posits because they motivated the use of mechanical principles, such as Hooke's law, in investigations of how light waves would propagate in such a medium. This led to the fundamental departure from previous wave theories brought about by Fresnel's assumption that light propagates as a transverse wave motion. Hence, since Fresnel got important 'heuristic mileage' out of these mechanical principles, then, as Worrall (1995) has argued, although the replacing electromagnetic theory provides a constructive proof of the eliminability of the old theory, mechanics is really only eliminable from the success of Fresnel's theory in a minimal logical sense.

When we inquire into what hypotheses 'really fuel the derivation' we have no other way to address this question than by explaining how we would derive the prediction in question using our understanding of the theory. This does not show that hypotheses which *we* do not use in reconstructing the derivation played no role in making it possible for the scientists of the time to derive the prediction that they did. A modern scientist may not need to invoke anything about the material constitution of the ether to reconstruct Fresnel's predictions, but Fresnel did so to derive the predictions in the first place. Psillos says that 'essential constituents' of success are ones such that 'no other available hypothesis' (Psillos 1999: 309) can replace them in the derivation of the novel predictions made by the theory in question. The question is: *available to whom?* There was no other hypothesis available to Fresnel about the nature of the ether that would have allowed him to derive the novel predictions of his theory. In general, it seems true that quite often there will be no other hypothesis available at the time, but that in reconstructing derivations we may have several alternatives.

Psillos argues that there is continuity between the causal roles of attributes of the ether and those of the field. For example, the fundamental causal role of the ether is arguably to act as the repository of

the energy associated with light between emission by a source and its absorption or reflection by matter. Light was known to travel at a finite velocity, so it had to be in some medium while passing through otherwise empty space. The electromagnetic field is now thought to be that medium. However, the selection of which causal role is important is done with the benefit of hindsight. Our assessment of what matters in the description of optical phenomena is very much relative to our current state of knowledge, as is any statement about the relevant causal role of some posited unobservable entity. However, we do not know what of current theories will be retained, in other words, what the *real* causal roles are.

It is true that the important principles about light (that it undergoes transverse propagation for example) are carried over into Maxwell's theory, and indeed there is a lot of continuity between the ether theories and electromagnetic field theory. However, the latter has now been replaced by quantum field theories, which may soon be replaced by a theory of superstrings or a grand unified theory of quantum gravity. It is implausible to suggest that 'ether' referred all along to a quantum field, because the latter has a completely different structure to Maxwell's electromagnetic field. For example, the latter is supposed to permeate all space and have a definite magnitude at different points, whereas the former is multi-dimensional and incorporates only probabilities for different magnitudes.

The case of caloric

In this case, Psillos argues that 'the approximate truth of a theory is distinct from the full reference of all its terms' (Psillos 1994: 161). Psillos argues that 'caloric' was not a *central* theoretical term, and we should only worry about the central theoretical terms: 'central in the sense that the advocates of the theory took the successes of the theory to warrant the claim that there are natural kinds denoted by these terms' (Psillos 1996: 312).

However, surely in the realism debate we ought only to be concerned with what scientists *should* believe and not with what they *do* in fact believe. If a group of antirealists came up with a novel predictively successful theory, realists would not admit that, because none of the scientists involved believes in the natural kinds

denoted by the theory, the theory has no central terms. When we are concerned with philosophical disputation about science the fact that a particular scientist had this or that attitude to a theory is irrelevant. If it happened that all the heat theorists had been instrumentalists we would not then deduce that we should be instrumentalists.

In any case, Psillos argues that in the case of caloric theory, all the important predictive successes of the theory were independent of the assumption that heat is a material substance. He argues that all the well-confirmed content of the theory – the laws of experimental calorimetry – was retained in thermodynamics anyway. What I think Psillos is doing here is exploiting, as scientists often do, the positive side of the underdetermination problem. When we theorise about or model the phenomena in some domain we will inevitably make a few mistakes. If there was only one (the true) theory that could describe the phenomena, then we would almost certainly never hit upon it first time, and we usually need to modify theories in the face of new evidence. More importantly, after radical theory change we want to be able to recover the empirical success of old theories without buying into their outdated ontologies. The ever present possibility of alternative, empirically equivalent theories is therefore essential.

The problem is that the only reason that we do not say that 'caloric' refers (to what we call heat or thermal energy) is because it is regarded as essential to this term that it is a material substance. However, Psillos says that the scientists of the time did not commit themselves to this, that they withheld judgement about the nature of caloric. Since the reason for saying that 'ether' refers is that scientists were not committed to it being of any particular nature but were committed to its having some properties that, as a matter of fact, are carried over to Maxwell's theory, we could say then that 'caloric' (like 'ether') does refer. On the other hand, suppose that scientists had been so committed to belief in the material constitution of the ether that denying its material nature would seem as absurd as denying that caloric is a (fictional) material substance. Then, by the same argument as that for caloric, we would say that 'ether' does not refer to anything, since the derivations of the important predictions of ether theory are all independent of

assumptions about its constitution (says Psillos), and that is the stated reason for saying that 'caloric' does not refer. Hence, whether we end up saying that this or that theoretical term refers to the same entities as a particular currently accepted theoretical term is arbitrary, and we could be in the position of saying that both 'ether' and 'caloric' did not refer.

It seems that even if Psillos' defence were to work, it would leave realism in a bit of a mess. Realists can no longer say that we should believe that the world is much like our best theories say it is, and that the theoretical terms of such theories genuinely refer. Instead, it is that some of these terms will refer, some not, and others will but only approximately. Furthermore, before realists commit themselves to any hypothesis or entity, they will have to look closely to see if it was essential in producing not just the success, but the novel predictive success of the theory. Where this is not the case they will not advise epistemic commitment. So, for example, before we know whether to believe that electrons are fundamental particles with spin 1/2 we had better check that their being spin 1/2 (and that their being fundamental and being particles) plays an essential role somewhere in the derivation of a novel prediction, and until this is done we had better be agnostic. On this account a theory can have great empirical success even though there is nothing in the world much like the entities postulated by the theory.

8.2 Multiple models

Roger Jones (1991) raises the problem 'Realism about what?'. He points to the existence of alternative formulations of physical theories that coexist in science, and argues that there is not necessarily a single picture of the world for us to believe in, even if we were assured of the epistemic warrant for doing so. For example, classical mechanics is a paradigm of a mature scientific theory and it has an impressive catalogue of predictive success from the return of Halley's comet to the recent successful space flights (classical mechanics is still used for such projects). The theory is first presented in terms of Newton's three laws of motion, understood as descriptive of a particle's

behaviour. Velocity and acceleration are introduced as functional relationships between a particle's positions at different times that are directly measurable. The theoretical concepts of mass and force are defined operationally: the ratio of the masses of two particles is a constant of proportionality that, together with their initial velocities and accelerations, determines the relative motions of the two particles after they have interacted; force is defined as simply the product of mass and acceleration. (This gives us a problem, for a particle that never interacts with anything else could have any value whatever for its mass, but since real particles will always interact with something or other let us ignore this.)

Classical mechanics is applied to planetary phenomena by combining the law of gravitation with the three laws of motion. This approach has a great deal of success in dealing with large approximately spherically symmetrical bodies whose mass can be conveniently regarded as zero everywhere except at the centre of mass. However, for extended bodies, in general it proves computationally intractable, and is replaced by the field theoretic formulation of classical mechanics. In this approach, the fundamental object is the gravitational potential, which is defined for all points of space and the magnitude and direction of the potential gradient at a point is what determines (with the mass) the magnitude and direction of the force experienced by matter at that point.

So far then, we have two different ways of formulating classical mechanics; either as a theory of forces acting at a distance between point particles, or as a theory describing a gravitational field that occupies all points of space and acts locally. There is a third formulation to consider, in which the force law and the laws of motion can be derived from so-called minimum principles. This is often known as 'analytic mechanics' and was first developed by Euler and Lagrange and given a full treatment by Hamilton in the nineteenth century. Analytic mechanics is derived by invoking the 'principle of least action', which states that the path followed by a particle between two points will be such as to minimise the so-called 'action integral', which represents the total energy, kinetic and potential, of the system. The last way to formulate classical mechanics, to which Jones draws attention, is the curved spacetime formulation analogous to that of general relativity. The gravitational field is absorbed into the

structure of spacetime and is represented by the degree of curvature at a given point.

Each of these approaches has similarities with other physical theories. The action at a distance approach was used for the formulation of the theory of the Coulomb force between charged particles in the late eighteenth century. The field-theoretic approach is directly analogous to the classical field theory of electromagnetism. Variational mechanics is related (mathematically at least) to general relativity and contemporary (gauge) field theories, and Hamiltonian mechanics is closely related to (non-relativistic) quantum mechanics. Suppose that our best theory of the world is classical mechanics and that matters of epistemology persuade us that we want to be realists; all this raises the question of what to be realistic about. The point is that each of these approaches suggests a quite different ontology and metaphysics. The first (Newtonian) approach suggests an ontology of point particles and forces acting at a distance. The field approach, on the other hand, suggests an ontology of action by contact in conformity with a principle of local causality, but postulates a new type of entity, the field. The curved spacetime approach suggests another picture again, where spacetime itself is part of the fundamental ontology of the theory and has a causal efficacy of its own. As for analytic mechanics, this seems to be independent of causal thinking, but requires instead a kind of teleology for particle behaviour according to which only properties of complete paths between points of space determine the motion of a body. The conclusion Jones draws from this is that the physicist, when asked to articulate the 'fundamental (theoretical) furniture of the Newtonian universe', will not be able to do so (Jones 1991: 190). According to Jones, classical mechanics is a 'vastly overconnected structure of concepts', which save the phenomena with very different explanatory frameworks and ontological commitments (Jones 1991: 190).

The short reply to Jones' problem would be just to point out that classical mechanics is known to be false, and that the existence of four alternative formulations of a false theory is no problem for the realist. However, this is an ultimately damaging response for the realist's position since it implies that they are only committed to a realist view of the one right theory of the world. Consider the situation at the end of the last century when Newtonian mechanics was widely

held, on the basis of great predictive and explanatory success, to be true. If realism has any content it must allow that people were justified in being realists to some extent about classical mechanics; the question that Jones poses is 'which version?'. Of course, the curved spacetime formulation was not available until early in the twentieth century, but the problem remains in principle.

Alan Musgrave replies to Jones' argument and the points he makes would be endorsed by many realists. Jones presents the four theories above as 'versions of the same theory', but Musgrave denies this:

> Can these empirically equivalent Newtonian theories simply be different 'versions' or 'formulations' of one and the same theory? We will think so if we accept the positivist or antirealist or verificationist principle that empirically equivalent theories are really the same theory.
>
> (Musgrave 1992: 693)

Therefore the suggestion is that they are not, in fact, the same theory on any but an antirealist construal along positivistic lines; even van Fraassen concedes that theories that refer to different entities and posit different explanatory frameworks (as Jones says these do) should be taken literally, and therefore really are different.

However, the reason Jones offers for calling these four versions of classical mechanics different formulations of the same theory is that they are all taught to students as part of classical mechanics. As Musgrave concedes, Jones' examples are of strong empirical equivalence and are more interesting than usual cases of underdetermination because they are 'embedded in the actual practice of scientists in a way that the artificial examples concocted by philosophers are not' (Musgrave 1992: 693). In any case, Musgrave claims, like most realists, that empirical equivalence does not entail evidential equivalence because explanatory power is evidentially significant. Of course, each of the four theories Jones cites does have explanatory power, so none of them can be disregarded as easily as can the 'The world is as if *T*' (call this a surrealist transform of *T*). However, Musgrave's point is that once we deny that empirical equivalence implies evidential equivalence, we will not immediately assume that there are no rational grounds for choosing between the four theories:

> Those who prefer causal to teleological explanation will rule out the minimum-principle 'version' of Newtonian mechanics. Does the history of science give us no reason to prefer causal to teleological explanation? Those who prefer local causality will rule out the action-at-a-distance 'version' of Newtonian mechanics. Scientists armed with a successful electromagnetic field theory might prefer the gravitational field theory because of the prospect it affords of unifying the two. And so forth.
>
> (Musgrave 1992: 696)

Although antirealists see this intrusion of metaphysics into physics as unwarranted, realists such as Musgrave believe that physics and metaphysics are continuous; each informs the other:

> Were Newton and many others foolish to worry about action-at-a-distance? Were Einstein and some others foolish to worry about the idealism that seemed implicit in the orthodox interpretation of quantum mechanics? Is it silly to oppose teaching our children creationist surrealism: 'God created the Universe in 4004 B.C. [or whenever] as if the teachings of natural science were true'?
>
> (Musgrave 1992: 696)

Musgrave wants to persuade us that the only way to rule out (what we may agree is) the madness of creation science, Berkeley's idealism and so on, is by using metaphysical considerations that will also rule out Jones' multiplicity problem. However, Berkeleyian idealism is rejected because it requires an ontology of ideas that is discredited for a whole host of reasons, none of which will help us with classical mechanics. Furthermore, as already pointed out, nobody *uses* surrealist transforms of theories, which is another reason for rejecting them that we cannot invoke to answer Jones. Creation science is ad hoc and has contributed no new empirical success to biology or palaeontology, and so it does not have anything like the credibility of evolutionary biology. Musgrave characterises his theory of evidential support as 'partly historical' (Musgrave 1992: 695) and argues that the prediction of 'novel facts' is evidentially important. However, none of the disreputable theories he cites have had any novel predictive success whereas all the formulations of classical mechanics can

claim some. It seems unwarranted to privilege the formulation of classical mechanics that happened to come first, when all of them could have predicted the same novel facts.

Even if we accept that explanatory power and metaphysical considerations are evidential, we have no easy solution to Jones' problem because different metaphysical criteria pull in different directions. Suppose that novel predictive success is our highest rated theoretical virtue; this would support the action-at-a-distance version of classical mechanics most. On the other hand, if we have a metaphysical penchant for local causality we will prefer the field theoretic version of the theory, which we will also prefer because of its compatibility with electromagnetic field theory. On the other hand, the variational approach to classical mechanics fits well with quantum mechanics, while the spacetime approach obviously coheres well with general relativity.

Musgrave ultimately claims: '[we] should be realists about what the best metaphysical considerations dictate, where the best metaphysical considerations are those that have yielded the best physics (Musgrave 1992: 691). However, there is no evidence that there is a unique set of metaphysical considerations that have yielded the best physics. On the contrary, the metaphysical presuppositions dominant at different stages in the history of science are themselves diverse and transient. Who would now argue that nature abhors a vacuum, or that all action must be contact between impenetrable particles, or indeed that a new theory cannot be adopted unless it admits of a mechanical reduction? Einstein was arguably influenced in the development of relativity theory by positivistic leanings, but surely this does not mean the realist should adopt positivism. There is also an apparent circularity in using metaphysical considerations to discover which is the best physics we have that then tells us what metaphysics to adopt.

8.3 Idealisation

The examples of scientific laws and theories used in philosophical debates about laws, explanation and so on are simplistic and perhaps misleading, given the chaos and multiplicity of the complex historical

processes of theory development and transmission. Moreover, since philosophers often do not discuss theories unless they have already come to form part of received scientific knowledge, the theories they discuss are generally the product of rational reconstruction by scientists. This has led to a critique of the methods and results of philosophers of science by those who have noticed the lack of correspondence between science as the object of philosophical analysis and science as a part of the lifeworld. Nancy Cartwright allies herself with those who have downplayed the significance of theory as opposed to practice for the philosophy of science, and advocates a new kind of instrumentalism:

> Our scientific understanding and its corresponding image of the world is encoded as much in our instruments, our mathematical techniques, our methods of approximation, the shape of our laboratories, and the pattern of industrial developments as in our scientific theories. [. . .] these bits of understanding so encoded should not be viewed as claims about the nature and structure of reality which ought to have a proper propositional expression that is a candidate for truth or falsehood. Rather they should be viewed as adaptable tools in a common scientific tool box.
>
> (Cartwright *et al.* 1995: 138)

Let us accept that there is much more to science than theories. Nonetheless, given that theories play some role in encoding our scientific knowledge, we may still inquire into what exactly that role is. Cartwright makes her view plain: 'Physics does aim to represent the world, but it represents it not in its theories but in its models' (Cartwright *et al.* 1995: 139).

One of Cartwright's central claims is that the traditional view of theories and models does not reflect the actual scientific practice of applying theories. The traditional view in question is the 'covering law' account: it suggests that theories need only to be conjoined with the appropriate auxiliary assumptions for particular contexts, and that together with these they will imply the data/phenomena. The actual use of theories and the construction of models in science involves the application of abstract theoretical structures to concrete situations. Cartwright's distinctive metaphysics of science is

motivated by her analysis of theory application. Her book, *How the Laws of Physics Lie*, made the distinction between *phenomenological* and *fundamental* laws and argued that the former refer to the concrete and the particular, and the latter to the abstract and the general. Furthermore: 'The phenomenological laws are indeed true of the objects in reality – or might be; but the fundamental laws are true only of the objects in the model' (Cartwright 1983: 4).

According to this book, the fundamental laws, because of their abstract nature, may be explanatory, but they do not describe what happens at all. They are linked to the appearances by phenomenological laws, which are non-explanatory but descriptive. Hence, at the theoretical level scientists construct models that are overtly of a sort that the real things do not fit. In order to relate these to specific phenomena they have to carry out a 'theory entry' process (Cartwright 1983: 132–134), whereby the phenomena are connected to theoretical models through a 'prepared description' that is overtly incorrect. This process is what we may loosely call 'idealisation'.

The most ubiquitous form of idealisation in physics is the application of mathematics to the physical world. For Duhem, this was enough to make the theoretical claims of physics quite separate from the ordinary truth-valued claims of everyday life, for the former are expressed in terms of concepts that are applied only with the help of artificially precise mathematics. Hence, he held that physical concepts are abstract and describe only imaginary constructions; they are mere symbolic formulae depending on a whole raft of theory for application to reality. Our concern here is not with 'mathematical idealisation', but with idealisations such as the point particle, the frictionless plane, the rigid rod and so on.

Cartwright distinguishes between cases where idealisation is performed on a concrete object/situation, and cases that are often called idealisation but where the simplifying assumptions involve abstracting, so that we are no longer dealing with concrete, but rather with abstract (and fictional), entities. For her, idealisation is the theoretical or experimental manipulation of concrete circumstances to minimise or eliminate certain features. For example, we idealise a real surface and get a frictionless plane, and then reintroduce a coefficient for friction with a convenient mathematical form. Here, the laws arrived at are approximately true, and in the laboratory it is possible to apply

them directly to very smooth surfaces and so on. Thus, the laws arrived at by idealisation are still *empirical* or *phenomenological*, that is they are still about *concrete* situations.

In the case of abstraction, we subtract concrete facts about objects, including perhaps the details of their material composition, and – especially importantly – we eliminate interfering causes. Hence, says Cartwright, the laws arrived at cannot be even approximately true, since relevant causal features have been subtracted and the laws are therefore not about concrete situations. In the traditional view, these abstract or fundamental laws are genuine claims about reality. Cartwright has argued that these laws lie if we interpret them as telling us about regularities in concrete situations, and that they are in fact *ceteris paribus* laws, and thus not true of any actual or possible concrete situation. For example, the law of gravitation states what happens to bodies upon which no other forces are acting; but there are no such bodies in the actual universe, and so strictly speaking it cannot be true of anything.

8.4 Structural realism

> [T]he structure of this physical world consistently moved farther and farther away from the world of sense and lost its former anthropomorphic character [. . .] Thus the physical world has become progressively more and more abstract; purely formal mathematical operations play a growing part.
>
> (Planck 1996: 41)

As we have seen, in the debate about scientific realism, arguably the two most compelling arguments around are the 'no miracles' argument, and the 'pessimistic meta-induction'. In an attempt to break this impasse and have 'the best of both worlds', John Worrall introduced *structural realism* (although he attributes its original formulation to Poincaré) (Worrall 1989). Using the case of the transition in nineteenth century optics from Fresnel's elastic solid ether theory to Maxwell's theory of the electromagnetic field, Worrall argues that:

> There was an important element of continuity in the shift from Fresnel to Maxwell – and this was much more than a simple

> question of carrying over the successful empirical content into the new theory. At the same time it was rather less than a carrying over of the full theoretical content or full theoretical mechanisms (even in 'approximate' form). [. . .] There was continuity or accumulation in the shift, but the continuity is one of form or structure, not of content.
>
> (Worrall 1989: 117)

According to Worrall, we should not accept full blown scientific realism, which asserts that the *nature* of things is correctly described by the metaphysical and physical content of our best theories. Rather, we should adopt the structural realist emphasis on the mathematical or *structural* content of our theories. Since there is (says Worrall) retention of structure across theory change, structural realism both (a) avoids the force of the pessimistic meta-induction (by not committing us to belief in the theory's description of the furniture of the world), and (b) does not make the success of science (especially the novel predictions of mature physical theories) seem miraculous (by committing us to the claim that the theory's *structure*, over and above its empirical content, describes the world).

Although structural realism has attracted considerable attention, Worrall's proposal needs to be developed if it is to provide a substantive alternative to traditional scientific realism. In particular, there is a fundamental question about the nature of structural realism that should be answered: is it metaphysics or epistemology? Worrall's paper is ambiguous in this respect. At times his proposal seems to be that we put an epistemic constraint on realism to the effect that we only commit ourselves to believing in the structural content of a theory, while remaining agnostic about the rest. This is suggested by Poincaré who talks of the redundant theories of the past capturing the 'true relations' between the 'real objects which Nature will hide for ever from our eyes' (Poincaré 1905: 161).

On the other hand, Worrall's position is not explicitly an epistemic one, and other comments suggest a departure from the metaphysics of standard scientific realism. For example, he says: 'On the structural realist view what Newton really discovered are the relationships between phenomena expressed in the mathematical equations of his theory' (Worrall 1989: 122). If the continuity in scientific change is of

'form or structure', then perhaps we should abandon commitment to the reference of theories to objects and properties, and account for the success of science in other terms. Redhead says: 'realism about what? Is it the entities, the abstract structural relations, the fundamental laws or what? My own view is that the best candidate for what is "true" about a physical theory is the abstract structural aspect' (Redhead 1996: 2). This seems to concur with structural realist sentiments expressed by Howard Stein: 'our science comes closest to comprehending "the real", not in its account of "substances" and their kinds, but in its account of the "Forms" which phenomena "imitate" (for "Forms" read "theoretical structures", for "imitate", "are represented by")' (Stein 1989: 57).

Structural realism has been the subject of recent debate in the philosophy of physics too but I cannot discuss it further here. I hope the reader will by now be aware that the question of scientific realism is much more complex than it first appears. In order to take the discussion further we would have to consider in detail the history of science, the nature of our best contemporary theories, and the implications of idealisation and the practice of science.

Further reading

On the meta-induction

Hardin, C.L. and Rosenberg, A. (1982) 'In defence of convergent realism', in *Philosophy of Science*, **49**, pp. 604–15.

Kitcher, P. (1993) *The Advancement of Science: Science without Legend, Objectivity without Illusions*, Oxford: Oxford University Press.

Laudan, L. (1981) 'A confrontation of convergent realism', *Philosophy of Science*, **48**, pp. 19–48, reprinted in D. Papineau (ed.) (1996) *Philosophy of Science*, Oxford: Oxford University Press.

Psillos, S. (1999) *Scientific Realism: How Science Tracks Truth*, London: Routledge.

Worrall, J. (1989) 'Structural realism: the best of both worlds?', *Dialectica*, **43**, pp. 99–124, reprinted in D. Papineau (ed.) (1996) *Philosophy of Science*, Oxford: Oxford University Press.

On reference

Field, H. (1995) 'Theory change and the indeterminacy of reference', in P. Lipton (ed.) *Theory, Evidence and Explanation*, Aldershot: Dartmouth.
McCulloch, G. (1989) *The Game of the Name: Introducing Logic, Language and the Mind*, Chapters 2 and 3, Oxford: Oxford University Press.
Papineau, D. (1979) *Theory and Meaning*, Section 5.6, Oxford: Oxford University Press.
Putnam, H. (1975) 'The meaning of "meaning"', in his *Mind, Language and Reality: Philisophical Papers*, Volume 2, Cambridge: Cambridge University Press.

Theories and models

Cartwright, N. (1983) *How the Laws of Physics Lie*, Oxford: Oxford University Press.
Dupré, J. (1993) *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Cambridge, MA: Harvard University Press.
Jones, R. (1991) 'Realism about what?', *Philosophy of Science*, **58**, pp. 185–202.
Musgrave, A. (1992) 'Discussion: realism about what?', *Philosophy of Science*, **59**, pp. 691–97.

Structural realism

Gower, B. (2000) 'Cassirer, Schlick and "structural" realism: the philosophy of the exact sciences in the background to early logical empiricism', *British Journal for the History of Science*, **8**, pp. 71–106.
Ladyman, J. (1998) 'What is structural realism?', *Studies in History and Philosophy of Science*.